

Summary of EPA Method Aquatic Life Criteria Development for Toxics

As directed by section 304(a) of the Clean Water Act, EPA develops and publishes national recommended criteria as guidance to states and tribes for the promulgation of their respective water quality standards. The law requires that these criteria be based on the latest scientific knowledge; no allowance is given to consider economics in criteria development.

For aquatic life, the national recommended toxics criteria are derived using a methodology published in *Guidelines for Deriving Numeric National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses*¹. Under these guidelines, criteria are developed from data quantifying the sensitivity of species to toxic compounds in controlled studies. The methodology allows for use of toxicity data from studies on aquatic animals and plants and from studies on the bio-concentration and bio-accumulation of toxins. However, almost all criteria are based on data quantifying the response of aquatic animals to toxins in laboratory tests.

Toxicity is expressed as the concentration of a toxin at which various effects occur, that is a given toxin can have more than one measure of its toxicity. Acute tests, performed with 48 or 96 hours of exposure, measure the concentration at which a toxin causes death in 50% of the individuals in the test, known as the LC50.

Chronic tests may last from weeks to years, depending on the lifespan of the species. Chronic tests are performed at sublethal toxin concentrations and measure effects on growth of individuals or their reproduction. Two effect levels are identified:

- 1) the concentration at which no observable effect occurs, known as the NOEC (e.g., no statistically significant reduction in growth)
- 2) the lowest concentration at which effects do occur, known as the LOEC (e.g., statistically significant reduction in growth)

The chronic criterion is derived from the geometric mean of these two effect levels.

The final recommended criteria are based on multiple species and toxicity tests. To develop a valid criterion, toxicity data must be available for at least one species in each of eight families of aquatic organisms. The eight taxa required are as follows:

- 1) salmonid (e.g., trout, salmon)
- 2) a fish other than a salmonid (e.g., bass, fathead minnow)
- 3) chordata (e.g., salamander, frog)
- 4) planktonic crustacean (e.g., daphnia)
- 5) benthic crustacean (e.g., crayfish)
- 6) insect (e.g., stonefly, mayfly)
- 7) rotifer, annelid (worm), or mollusk (e.g., mussel, snail)

¹ Stephan, C.E. and others. 1985. *Guidelines for Deriving Numeric National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses*. US Environmental Protection Agency, Office of Research and Development, Environmental Research Laboratories, Duluth, MN, Narragansett, RI, Corvallis, OR. EPA 822/R-85-100.

8) a second insect or mollusk not already represented above

Where toxicity data are available for multiple life stages of the same species (e.g., eggs, juveniles, and adults), the procedure requires that the data from the most sensitive life stage be used for that species. This assures individuals of that species can survive an exposure at any time during their life, and thus has the opportunity to produce a succeeding generation.

Toxicity tests are screened for validity; screening criteria include:

- ✓ use of a control
- ✓ testing a of single species and compound per test
- ✓ species tested is from North America
- ✓ flow through test used if compound is volatile or easily degrades
- ✓ hardness or pH reported when relevant to toxicity

All valid toxicity data are used. If there are multiple tests for a species, which is often the case, the data are pooled and averaged to provide a species mean toxic concentration. If data from several species in a single genus are available, the species mean values are pooled to calculate a genus mean toxic concentration. If data exist from only one species, that species mean value becomes the genus mean value.

The genus mean values, from the minimum eight families of aquatic organisms, are then ranked from high (most resistant to the toxin) to low (most sensitive to the toxin). The lowest four values are then used in regression to estimate the concentration that would cause the threshold effect (i.e., LC50) for the fifth percentile most sensitive species. Since this fifth percentile is statistically derived, it is unlikely to correspond to any of the genus mean values actually calculated.

For acute toxicity tests, the fifth percentile of the effect concentrations is known as the final acute value (FAV). If a *species* mean value for a commercially or recreationally important species, such as rainbow trout, falls below the FAV, the guidelines say that this species mean value can be substituted for the fifth percentile of the *genus* mean values to protect that important species.

The FAV is divided by two (a safety factor to avoid lethality) to arrive at the recommended acute criterion. As an additional safety factor, the acute criterion is applied as a limit on 1-hour average concentration in the environment, though in testing the exposure is for duration of 48 or 96 hours.

Water hardness (the amount of calcium and magnesium in the water) is known to moderate the toxicity of some metals; in particular, the toxicity of cadmium (Cd), chromium III (Cr III), copper (Cu), lead (Pb), nickel (Ni), silver (Ag), and zinc (Zn) decrease as water hardness increases. For these toxins, a relationship of toxicity to hardness is calculated first. Because their toxicity is dependent on water hardness, the effect concentrations from toxicity tests at differing hardness levels are adjusted (normalized) to a hardness of 50 milligrams per liter. Then, the process of deriving the FAV is begun. To derive the toxicity-versus-hardness relationship for a given metal, only toxicity studies where hardness was varied are useful, which is only a small

subset of all test results. When deriving the FAV for these metals, a study must have measured and reported water hardness to be valid, even if the values are not varied in the study. The resulting criteria are hardness-dependent equations, not fixed values.

The same procedure is preferred for chronic test results. However, chronic toxicity tests are much more expensive and, therefore, not as common as acute tests. Usually, there is not enough chronic toxicity testing data to meet the minimum requirement of eight families of aquatic organisms. For this situation the guidelines provide an alternate way to derive a chronic criterion from the FAV using ratios derived from studies in which both acute and chronic tests have been conducted simultaneously for the same species.

Acute-chronic ratios are calculated for each set of parallel tests, then averaged (geometric mean) to arrive at the final acute-chronic ratio. Three studies with parallel testing are required to calculate a valid final ratio. The chronic criterion is then calculated from the FAV (not the acute criterion) by dividing it by the final acute-chronic ratio.

So, if the side-by-side tests show that, on average, chronic effects occur at one-half the concentration of acute effects, the chronic criterion will be one-half of the FAV, which is the same as the acute criterion. Typically, chronic effects show up at even lower concentrations, so the chronic criterion is usually lower than the acute criterion. Chronic criteria are applied as a limit on the four-day average concentration in the environment. Both the acute and chronic criteria are values that are not to be exceeded more than once in three years. In other words, the criteria specify a magnitude, duration, and frequency to be met in order to provide protection of aquatic life.